

AD-A145 579

ON DISCRETE FAILURE MODELS(U) SOUTH CAROLINA UNIV

1/1

COLUMBIA DEPT OF MATHEMATICS AND STATISTICS

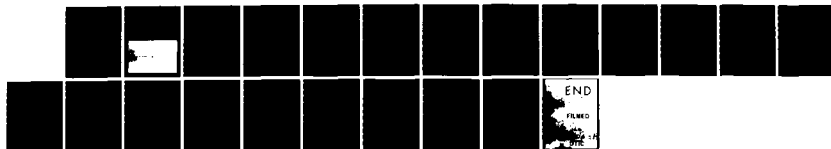
W J PADGETT ET AL. 17 JUL 84 TR-97 AFOSR-TR-84-0667

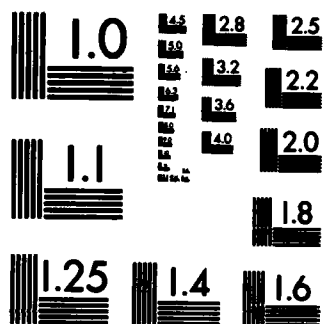
UNCLASSIFIED

AFOSR-84-0156

F/G 14/2

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

Department of Mathematics  
and Statistics  
The University of South Carolina  
Columbia, South Carolina 29208

AFOSR-TR- 84 08 17 103

17 1984

A  
4

DTIC FILE COPY

AD-A145 579

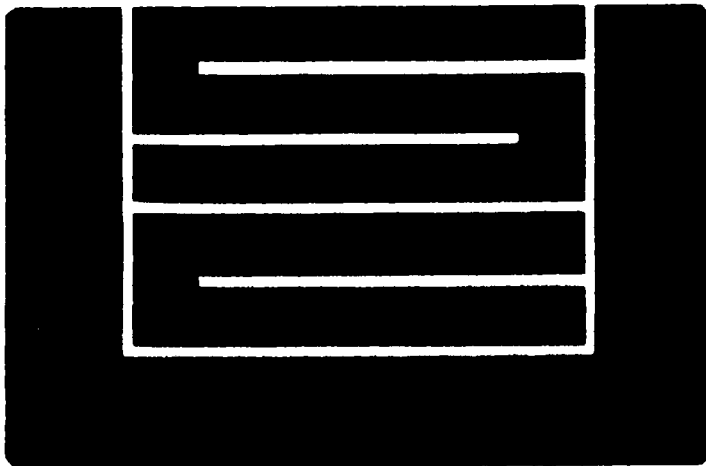
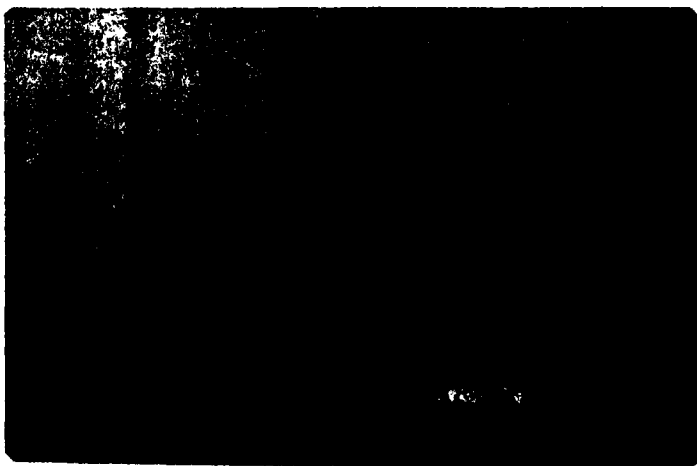
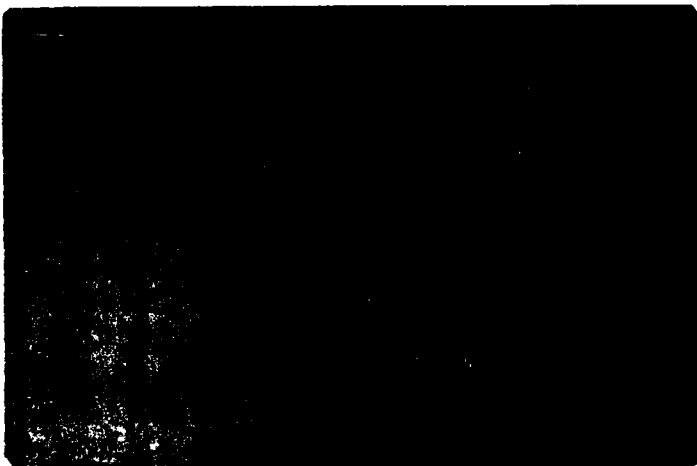
ON DISCRETE FAILURE MODELS\*

by

W. J. Padgett and John D. Spurrier

University of South Carolina  
Statistics Technical Report No. 97  
62N05-10

Approved for  
distribution



H

ON DISCRETE FAILURE MODELS\*

by

W. J. Padgett and John D. Spurrier

University of South Carolina  
Statistics Technical Report No. 97  
62N05-10

July, 1984

Department of Mathematics and Statistics  
University of South Carolina  
Columbia, South Carolina 29208

1984  
A

Chief, Technical Information Division

\* Research supported by the United States Air Force Office of Scientific Research and Army Research Office under Grant No. AFOSR-84-0156.

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1d. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			5. MONITORING ORGANIZATION REPORT NUMBER(S) <b>AFOSR-TR- 84 - 0 6 6 7</b>		
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Stat. Tech. Rep. No. 97 (62N05-10)			7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research		
6a. NAME OF PERFORMING ORGANIZATION University of South Carolina		6b. OFFICE SYMBOL (if applicable)	7b. ADDRESS (City, State, and ZIP Code) Directorate of Mathematical & Information Sciences, AFOSR, Bolling AFB DC 20332		
6c. ADDRESS (City, State, and ZIP Code) Department of Mathematics & Statistics Columbia SC 29208		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-84-0156			
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR		8b. OFFICE SYMBOL (if applicable) M1	10. SOURCE OF FUNDING NUMBERS		
8c. ADDRESS (City, State, and ZIP Code) Bolling AFB DC 20332		PROGRAM ELEMENT NO. 61102F	PROJECT NO. 2304	TASK NO. A5	WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) ON DISCRETE FAILURE MODELS.					
12. PERSONAL AUTHOR(S) W.J. Padgett and John D. Spurrier					
13a. TYPE OF REPORT Technical		13b. TIME COVERED FROM TO		14. DATE OF REPORT (Year, Month, Day) 17 JUL 84	
15. PAGE COUNT 18					
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Discrete hazard functions; increasing failure rate; decreasing failure rate; discrete life data; maximum likelihood estimation.		
19. ABSTRACT (Continue on reverse if necessary and identify by block number) In some situations, discrete failure time distributions are more appropriate to model life-times than continuous distributions. Very few results for the discrete case appear in the literature. This paper provides three families of discrete parametric lifetime distributions which are quite versatile in fitting IFR, DFR, and constant failure rate models to either uncensored or right-censored life-test data. The maximum likelihood estimation of parameters survival probabilities, and mean lifetimes is investigated, and the MLEs are shown to be easily computed by simple numerical methods. An example is given for each of the models, allowing the comparison of the proposed models. The example illustrates that the discrete models presented here can provide a better fit to discrete data than previously proposed discrete distributions.					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a. NAME OF RESPONSIBLE INDIVIDUAL CPT Brian W. Woodruff			22b. TELEPHONE (Include Area Code) (301) 767- 5027		22c. OFFICE SYMBOL

## ON DISCRETE FAILURE MODELS

W. J. Padgett and John D. Spurrier

University of South Carolina, Columbia

Key Words - Discrete hazard functions; Increasing failure rate;  
Decreasing failure rate; Discrete life data; Maximum likelihood estimation.

Summary & Conclusions: In some situations, discrete failure time distributions are more appropriate to model lifetimes than continuous distributions. Very few results for the discrete case have been given in the literature. This paper provides three families of discrete parametric life-time distributions which are quite versatile in fitting <sup>(Increasing Failure Rate)</sup> (IFR), <sup>(Decreasing Failure Rate)</sup> (DFR), and constant failure rate models to either uncensored or right-censored life-test data. The maximum likelihood estimation <sup>(MLE)</sup> of parameters, survival probabilities, and mean lifetimes is investigated, and the MLEs are shown to be easily computed by simple numerical methods. An example is given for each of the models, allowing the comparison of the proposed models. The example illustrates that the discrete models presented here can provide a better fit to discrete data than previously proposed discrete distributions. A



## 1. INTRODUCTION

A great amount of research has been performed for continuous life distributions. However, very little has appeared in the literature for discrete failure models. Discrete failure data arise in several common situations: (i) A device can be monitored only once per time period (an hour, a day, etc.), and the observation taken is the number of time periods successfully completed prior to the failure of the device. (ii) A piece of \_\_\_\_\_ equipment may operate in cycles, and the experimenter observes the number of cycles successfully completed prior to failure. Many other examples may be cited.

In situations where the observed data values are very large, in thousands of cycles, etc., a continuous distribution can be an adequate model for the discrete random variable. For example, the Birnbaum-Saunders [3] distribution arises in connection with fatigue failure due to bending of metal alloys. However, when the observed values are small, a continuous distribution may not adequately describe a discrete random variable, and an appropriate discrete failure model is desirable.

A few results on discrete life distributions have appeared in the literature. Kalbfleisch & Prentice [5, Ch. 2] briefly discuss discrete failure distributions that may be obtained from well-known continuous distributions such as the exponential or Weibull. Langberg, Leon, Lynch & Proschan [5,6] examine properties of the classes of discrete DFR and DFRA distributions. Salvia & Bollinger [9] consider the properties of discrete hazard functions and analogies with the continuous case. Barlow, Bartholomew, Bremner & Brunk [1] give a nonparametric estimate of an IFR function using isotonic regression

techniques. Their estimate is easily modified for DFR functions and to allow for censoring.

The purpose of this paper is to further examine discrete hazard functions in a parametric setting. Estimation for the discrete hazard functions given by Salvia & Bollinger [9] will be discussed for both uncensored and censored lifetime data in Section 2. In addition, some other parametric discrete failure models will be proposed and estimation for them will be studied in Sections 3 and 4. It should be noted that meaningful discrete parametric models generally present difficult estimation problems and may not be founded on a theoretical basis [5, p. 36]. However, such models are still quite important for discrete failure data.

#### Notation

MLE	Maximum Likelihood Estimator
IFR	Increasing Failure Rate
DFR(A)	Decreasing Failure Rate (Average)
$f_k$	$pmf\{k\} = \Pr\{X = k\}$
$S_k$	$Sf\{k\} = \Pr\{X \geq k\}$
$h_k$	discrete hazard, $f_k/S_k$
$k_1, \dots, k_n$	random sample of $n$ observations on a discrete random variable $X$
$\alpha, \hat{\alpha}$	a parameter and its MLE
$F$	a cumulative distribution function
$E(X)$	expectation of the random variable $X$
$L, \ln L$	a likelihood function and its natural logarithm

## 2. THE MODELS OF SALVIA AND BOLLINGER

Salvia & Bollinger [9] developed some results for discrete failure models analogous to the theory of the continuous case. Some of their results follow.



The discrete hazard rate (or failure rate) is related to the pmf by

$$h_k = f_k / (f_k + f_{k+1} + \dots) \text{ and } f_0 = h_0, f_1 = h_1(1-h_0), \dots,$$

$$f_k = h_k(1-h_0)(1-h_1) \dots (1-h_{k-1}). \text{ Similarly the discrete Sf is}$$

$$S_k = (1-h_0)(1-h_1) \dots (1-h_{k-1}) \text{ for } k = 1, 2, \dots. \text{ The mean life, if it exists,}$$

$$\text{is } E(X) = \sum_{k=1}^{\infty} S_k = \sum_{k=1}^{\infty} \prod_{j=0}^{k-1} (1-h_j). \text{ It should be noted that using Fort}$$

[4, p. 32-33],  $E(X)$  exists if there exists an  $\epsilon > 0$  and an  $m$  such that

for  $k > m$ ,  $k[h_k/(1-h_k)] > 1 + \epsilon$ . If for  $k > m$ ,  $k[h_k/(1-h_k)] \leq 1$ , then

$E(X)$  does not exist. Some other results analogous to continuous failure models

were also given in [9] for certain situations. For example, for a distribution with

IFR,  $h_0 \leq h_1 \leq h_2 \leq \dots$ ,  $S_k \leq (1-h_0)^k$  and  $E(X) \leq (1-h_0)/h_0$ . These inequalities

are reversed for DFR distributions.

Salvia & Bollinger [9] also gave illustrations of discrete parametric failure models for the constant failure rate, IFR, and DFR cases.

a. Constant Failure Rate. For  $h_k = c$ , a constant  $0 < c \leq 1$  for all  $k=0, 1, 2, \dots$ ,  $f_k = c(1-c)^k$ , a geometric distribution which is the discrete analog of the exponential distribution. We will give the MLE of  $c$  for the three cases of a complete random sample  $k_1, \dots, k_n$  from  $f_k$ , a type II censored sample, and a type I censored sample.

For a complete random sample, the log-likelihood is  $\ln L(c; k_1, \dots, k_n) = n \ln c + \sum_{i=1}^n k_i \ln(1-c)$ , which is uniquely maximized by  $\hat{c} = n / (n + \sum_{i=1}^n k_i)$ . Hence, the MLE of  $S_k$  is  $\hat{S}_k = (1-\hat{c})^k$ ,  $k=0, 1, 2, \dots$ .

For a type II censored sample, denote  $k_{(1)} = \min\{k_1, \dots, k_n\}$ ,  $k_{(2)}$  = next smallest of  $\{k_1, \dots, k_n\}$ , ...,  $k_{(r)}$  =  $r$ th smallest of  $\{k_1, \dots, k_n\}$ , where some of the  $k_i$ 's may be equal. If  $r+m$  items fail at or before  $k_{(r)}$ ,  $m=0, 1, \dots, n-r$ , the log-likelihood is

$$\ln L(c; k_1, \dots, k_n) = (r+m) \ln c + \left( \sum_{i=1}^{r+m} k_{(i)} + (n-r-m)(k_{(r)} + 1) \right) \ln(1-c)$$

which is uniquely maximized by

$$\hat{c} = (r+m) / \left[ (r+m) + \sum_{i=1}^{r+m} k_{(i)} + (n-r-m)(k_{(r)} + 1) \right]. \quad (2.1)$$

For a type I censored sample, where the life test is terminated after a fixed number of "cycles" or time units  $T$ , the log-likelihood function is  $\ln L(c; k_1, \dots, k_n) = \sum_{i=1}^r \ln f_{k_{(i)}} + (n-r) \ln S_T$ , where  $r$  is the observed number of failures by the end of  $T$  cycles. The MLE  $\hat{c}$  for this case is similar to (2.1) and is omitted. In fact, the expressions in  $\ln L$  for arbitrarily right-censored data, which includes type I and type II censoring, have the same general form for this or any of the discrete models which follow.

b. An IFR Distribution. For  $h_k = 1 - c/(k+1)$ ,  $k=0,1,2,\dots$ ,  $0 \leq c \leq 1$ ,

$f_k = (k-c+1)c^k/(k+1)!$ , an IFR distribution.

c. A DFR Distribution. For  $h_k = c/(k+1)$ ,  $k=0,1,2,\dots$ ,  $0 < c \leq 1$ ,

$f_k = c(1-c)(2-c)\dots(k-c)/(k+1)!$ , a DFR distribution.

We have found that these IFR and DFR distributions are not sufficiently flexible to fit a wide variety of settings. For example, for the IFR distribution the MLE  $\hat{c} = 1$  if all  $k_i$ 's  $> 0$ , regardless of the values of the  $k_i$ 's. For the DFR distribution  $E(X)$  exists only for  $c=1$ . Generalizations of these distributions and estimation results are given in Section 4.1.

### 3. DISCRETE FAILURE MODELS FROM CONTINUOUS DISTRIBUTIONS

Kalbfleisch & Prentice [5, pp. 35-36] indicate that failure time data are sometimes discrete either because of the grouping of continuous data due to

imprecise measurement or because time itself is discrete. They also give discrete distributions based on grouping data from continuous distributions into unit intervals. For example, if we begin with the Weibull distribution  $\text{weif}(\beta t; \alpha) = 1 - \exp(-(\beta t)^\alpha)$ ,  $t \geq 0$ , then the discrete distribution obtained by setting  $f_k = \text{weif}(\beta(k+1); \alpha) - \text{weif}(\beta k; \alpha) = \exp(-\beta k)^\alpha - \exp(-\beta(k+1))^\alpha$ ,  $k=0,1,2,\dots$ , is a distribution for grouped failure data. If  $\alpha = 1$ ,  $f_k$  becomes  $f_k = \exp(-\beta k)(1 - \exp(-\beta)) = c(1-c)^k$ ,  $k=0,1,2,\dots$ , where  $c = 1 - \exp(-\beta)$ , which is the geometric model (a) of Salvia & Bollinger [9]. If  $\alpha < 1$ , a discrete DFR distribution is obtained, and if  $\alpha > 1$ ,  $f_k$  is an IFR distribution.

More generally, it is easy to show that if a continuous IFR(DFR) distribution is used to generate a discrete distribution  $f_k$ , then  $f_k$  is also IFR(DFR). However, the class of continuous distributions giving rise to discrete IFR, DFR, or constant failure rate distributions is not restricted to being IFR, DFR, or constant failure rate, respectively. For example, the pdf  $g(x) = \beta \exp(-\beta x) + 2 \exp(-\beta(k+\frac{1}{2})) - 2 \exp(-\beta(k+1))$ ,  $k \leq x \leq k + \frac{1}{2}$ , and  $g(x) = 0$ ,  $k + \frac{1}{2} < x < k+1$ ,  $k=0,1,2,\dots$  (and  $g(x) = 0$ ,  $x < 0$ ) does not have constant failure rate, IFR, nor DFR, yet  $f_k = \exp(-\beta k)(1 - \exp(-\beta))$ ,  $k=0,1,2,\dots$ , which has constant failure rate, as before.

#### 4. SOME PROPOSED DISCRETE PARAMETRIC FAILURE MODELS

We propose some more general discrete failure models than those of Salvia & Bollinger [9]. First, a generalization of their IFR and DFR models which were mentioned in Section 2 is given and maximum likelihood estimation of the parameters is investigated. Then a discrete failure model which can have increasing, decreasing, or constant failure rate according to the values of one of its parameters is presented, and estimation of the parameters is considered.

It should be noted also that some of the well-known distributions, in addition to the geometric distribution, may be useful in modeling discrete life data. For example, it may be easily shown directly, or using the results on pp. 76-77 of Barlow & Proschan [2], that the Poisson distribution is an IFR distribution.

#### 4.1 Generalized Salvia and Bollinger Models

The IFR and DFR models of Salvia & Bollinger [9] given in Section 2 will be generalized by the addition of a second parameter  $\alpha$ . When  $\alpha=1$  the models reduce to those in Section 2. When  $\alpha=0$  the models reduce to the constant failure rate model of Section 2.

##### The IFR Case (Model 1)

Let  $h_k = 1 - c/(\alpha k + 1)$ ,  $k=0,1,2,\dots$ , ( $0 \leq c \leq 1$ ,  $0 \leq \alpha$ ). The pmf is  $f_k = (\alpha k + 1 - c)c^k / \prod_{j=0}^k (\alpha j + 1)$ ,  $k=0,1,2,\dots$ , and the Sf is  $S_k = c^k / \prod_{j=0}^{k-1} (\alpha j + 1)$ ,  $k=1,2,\dots$ . As  $\alpha \rightarrow \infty$ , the model approaches the Bernoulli model with  $f_0 = 1 - c$  and  $f_1 = c$ . As  $c \rightarrow 0$ , the model approaches the degenerate distribution with  $f_0 = 1$ . The parameter  $c = P_r\{X > 0\}$  and  $\alpha$  is a shape parameter. The pmf is nonincreasing in  $k$  if  $c > \frac{1}{2}$  and  $\alpha \leq (c-1)^2/(2c-1)$  or if  $c \leq \frac{1}{2}$ . Otherwise, the pmf increases and then decreases in  $k$ . Figure 1 shows the pmf's for some values of  $\alpha$  and  $c$ .

We now consider the estimation of the parameters from an observed random sample  $k_1, \dots, k_n$  from  $f_k$ . The log-likelihood function is

$$\ln L = \sum_{i=1}^n k_i \ln c + \sum_{i=1}^n \ln(\alpha k_i + 1 - c) - \sum_{i=1}^n \sum_{j=0}^{k_i} \ln(\alpha j + 1). \quad (4.1.1)$$

Three cases will be considered.

Case i:  $\alpha$  Known. If all  $k_1 = 0$ , then  $\ln L = n \ln(1-c)$ , which gives  $\hat{c} = 0$ . Next, if at least one  $k_1 > 0$ , from (4.1.1)

$$\frac{\partial}{\partial c} \ln L = \sum_{i=1}^n k_i / c - \sum_{i=1}^n 1 / (\alpha k_i + 1 - c). \quad (4.1.2)$$

Then  $\frac{\partial^2}{\partial c^2} \ln L < 0$  for all  $0 < c < 1$ . It follows from examining the limits of (4.1.2) as  $c \rightarrow 0$  and  $c \rightarrow 1$  that there is a unique value  $\hat{c}$  which maximizes (4.1.1). If all  $k_1 > 0$  and  $\alpha \leq \frac{\sum_{i=1}^n (1/k_i)}{\sum_{i=1}^n k_i}$ , then  $\hat{c} = 1$ . Otherwise,  $\hat{c}$  may be approximated by solving  $\frac{\partial}{\partial c} \ln L = 0$  for  $c$  using the Newton-Raphson procedure.

Case ii:  $c$  Known. If all  $k_1$ 's = 0, then (4.1.1) does not involve  $\alpha$ . Hence,  $\hat{\alpha}$  is arbitrary. If at least one  $k_1 > 0$ , then  $\frac{\partial}{\partial \alpha} \ln L = 0$  may or may not have a solution for  $\alpha$ .

Case iii: Both  $\alpha$  and  $c$  Unknown. If all  $k_1$ 's = 0, then  $\hat{c} = 0$  and  $\hat{\alpha}$  is arbitrary. Next, if  $n_0$   $k_1$ 's = 0 and the remaining  $k_1$ 's = 1, then  $\hat{c} = (n - n_0)/n$  and  $\hat{\alpha} = +\infty$ . Finally, if at least one  $k_1 > 1$ , it is difficult, if not impossible to show that unique MLE's always exist. However, it is possible to find the unique value of  $c$  which maximizes (4.1.1) for fixed  $\alpha$ . A line search of the variable  $\alpha$  can then be performed to obtain an approximate maximum for (4.1.1).

For the cases of type I and type II censoring, the MLE's for this model are obtained in the same manner. For type I censoring after  $T$  cycles, the log-likelihood function is

$$\begin{aligned} \ln L = & \ln(c) [(n-r) + \sum_{i=1}^r k_{(i)}] + \sum_{i=1}^r \ln [\alpha k_{(i)} + 1 - c] \\ & - \sum_{i=1}^r \sum_{j=0}^{k_{(i)}} \ln(\alpha j + 1) - (n-r) \sum_{j=0}^{T-1} \ln(\alpha j + 1). \end{aligned} \quad (4.1.3)$$

For type II censoring with  $r+m$  failures at or before  $k_{(r)}$ ,  $m=0,1,\dots,n-r$ , the log-likelihood is

$$\begin{aligned} \ln L = \ln(c) [(n-m-r)(k_{(r)}+1) + \sum_{i=1}^{m+r} k_{(i)}] + \sum_{i=1}^{m+r} \ln(\alpha k_{(i)}+1-c) \\ - \sum_{i=1}^{m+r} \sum_{j=0}^{k_{(i)}} \ln(j\alpha+1) - (n-m-r) \sum_{j=0}^{k_{(r)}} \ln(j\alpha+1). \end{aligned} \quad (4.1.4)$$

The derivatives of (4.1.3) and (4.1.4) have the same behavior as those of (4.1.1), so the MLE's may be obtained numerically as described for uncensored data.

The mean lifetime for this model from the formula in Section 2 is

$$E(X) = \sum_{k=1}^{\infty} c^k / \prod_{j=0}^{k-1} (j\alpha+1) \text{ which may be easily calculated to}$$

any desired degree of accuracy by summing an appropriate number of terms of the series. An estimate of  $E(X)$  is found by replacing  $\alpha$  and  $c$  by  $\hat{\alpha}$  and  $\hat{c}$ .

#### The DFR Case (Model 2)

Let  $h_k = c/(\alpha k+1)$ ,  $k=0,1,2,\dots$ , ( $0 < c \leq 1$ ,  $0 \leq \alpha$ ). The pmf is  $f_0 = c$  and

$$f_k = c \prod_{j=0}^{k-1} (j\alpha+1-c) / \prod_{j=0}^{k-1} (j\alpha+1), \quad k=1,2,\dots,$$

and the Sf is  $S_k = \prod_{j=0}^{k-1} (j\alpha+1-c)/(j\alpha+1)$ , where  $\prod_{j=0}^{-1} \equiv 1$ . As  $c \rightarrow 1$ , the model approaches the degenerate distribution with  $f_0 = 1$ . The parameter  $c = \Pr\{X=0\}$  and  $\alpha$  is a shape parameter. The pmf is decreasing in  $k$  for all  $\alpha$  and  $c$ . Figure 2 shows the pmf's for various choices of  $\alpha$  and  $c$ .

We now consider the estimation of the parameters from an observed random sample  $k_1, \dots, k_n$  from  $f_k$ . The log-likelihood function is

$$\ln L = n \ln c + \sum_{i=1}^n \sum_{j=0}^{k_i-1} \ln(j\alpha+1-c) - \sum_{i=1}^n \sum_{j=0}^{k_i-1} \ln(j\alpha+1), \quad (4.1.5)$$

where  $\sum_{j=0}^{-1} \equiv 0$ .

The results for maximizing (4.1.5) are analogous to the IFR case. If  $\alpha$  is known, there is a unique  $0 < \hat{c}_\alpha \leq 1$  which maximizes (4.1.5). This  $\hat{c}_\alpha$  can be approximated by the Newton-Raphson procedure. If  $c$  is known, then  $\frac{\partial}{\partial \alpha} \ln L = 0$  may or may not have a solution. If both  $\alpha$  and  $c$  are unknown, then an approximate maximum for (4.1.5) is found by performing a line search on  $\alpha$  with  $c = \hat{c}_\alpha$ .

For the cases of type I and type-II censoring the MLE's are obtained in the same manner. For type I censoring after  $T$  cycles, the log-likelihood function is

$$\begin{aligned} \ln L = r \ln c + \sum_{i=1}^r \left[ \sum_{j=0}^{k(i)-1} \ln(j\alpha+1-c) - \sum_{j=0}^{k(i)} \ln(j\alpha+1) \right] \\ + (n-r) \sum_{j=0}^{T-1} \ln[(\alpha j+1-c)/(\alpha j+1)]. \end{aligned} \quad (4.1.6)$$

For type II censoring with  $r+m$  failures at or before  $k_{(r)}$ ,  $m=0,1,\dots,n-r$ , the log-likelihood is

$$\begin{aligned} \ln L = (m+r) \ln c + \sum_{i=1}^{m+r} \left[ \sum_{j=0}^{k(i)-1} \ln(j\alpha+1-c) - \sum_{j=0}^{k(i)} \ln(j\alpha+1) \right] \\ + (n-m-r) \sum_{j=0}^{k(r)} \ln[(\alpha j+1-c) / (\alpha j+1)]. \end{aligned} \quad (4.1.7)$$

The MLEs may be obtained numerically as for uncensored data.

Using the results in Section 2,  $E(X)$  exists if  $c=1$  or  $1 > c > \alpha$  and does not exist if  $c \leq \alpha$  and  $c \neq 1$ . If  $c > \alpha$ , then the mean lifetime may be easily approximated by summing an appropriate number of terms of

$$E(X) = \sum_{k=1}^{\infty} \prod_{j=0}^{k-1} (\alpha j+1-c) / (\alpha j+1).$$

An estimate of  $E(X)$  is found by replacing  $\alpha$  and  $c$  by  $\hat{\alpha}$  and  $\hat{c}$ .

#### 4.2 A Discrete "Weibull Hazard" Model (Model 3)

Let  $h_k = 1 - \exp(-c(k+1)^\alpha)$ ,  $k=0,1,2,\dots$ , ( $c > 0$ ,  $-\infty < \alpha < \infty$ ). Since this  $h_k$  is similar to the Weibull Cdf, we will call this the discrete Weibull hazard model, or Model 3. Thus the pmf is

$$f_k = (1 - \exp(-c(k+1)^\alpha)) \prod_{i=1}^k \exp(-ci^\alpha) \\ = \exp(-c \sum_{i=1}^k i^\alpha) (1 - \exp(-c(k+1)^\alpha)), \quad k=0,1,2,\dots,$$

and the Sf is  $S_k = \exp(-c \sum_{i=1}^k i^\alpha)$ ,  $k=1,2,\dots$ , where  $\sum_{i=1}^0 \equiv 0$ .

Note that for  $\alpha=0$ , this model reduces to the constant failure rate model of Section 2. For  $\alpha > 0$ , the distribution has IFR, and for  $\alpha < 0$ , it has DFR. Thus, this discrete model is quite flexible with respect to choice of failure rate, analogous to the Weibull distribution in the continuous case. The parameter  $\alpha$  is essentially a shape parameter and  $c$  is a "spread" parameter. For small values of  $c$ , the probability is spread out in the tail of the distribution, whereas for large  $c$ , most of the probability is at  $k=0$ . For example, for  $c=5$ ,  $f_0 = 0.9933$ . For  $\alpha < 0$ , the pmf is always decreasing in  $k$ . For  $\alpha > 0$ , the pmf is either decreasing or first increasing and then decreasing in  $k$ . In particular, for small  $c$ , the largest  $f_k$  is at large values of  $k$ . For example, for  $\alpha = 0.5$ ,  $c = .01$ , the largest value of  $f_k$  is at  $k=13$ . Figure 3 shows the pmf's for some values of  $\alpha$  and  $c$ . It appears that for various values of  $c > 0$ ,  $\alpha$  in the range  $-1 \leq \alpha \leq 1$  is sufficient to describe many IFR, DFR, or constant failure rate discrete distributions.

We now consider the estimation of the parameters from an observed random sample  $k_1, \dots, k_n$  from  $f_k$ . The log-likelihood function is

$$\ln L = \sum_{i=1}^n \ln[1 - \exp(-c(k_i+1)^\alpha)] - c \sum_{i=1}^n \sum_{j=1}^{k_i} j^\alpha. \quad (4.2.1)$$



Three cases will be considered.

Case i:  $\alpha$  Known. If at least one  $k_1 > 0$ , from (4.2.1)

$$\frac{\partial}{\partial c} \ln L = \sum_{i=1}^n (k_i + 1)^\alpha / [\exp(c(k_i + 1)^\alpha) - 1] - \sum_{i=1}^n \sum_{j=1}^{k_i} j^\alpha. \quad (4.2.2)$$

Then,  $\frac{\partial^2}{\partial c^2} \ln L < 0$  for all  $c > 0$ , and as  $c \rightarrow 0$ , (4.2.2) approaches  $+\infty$ , whereas as  $c \rightarrow \infty$ , (4.2.2) approaches a negative constant. Hence, there is a unique MLE for  $c$ . This MLE can be found by solving  $\frac{\partial}{\partial c} \ln L = 0$  numerically using Newton-Raphson iteration.

If  $k_1 = 0$  for all  $i=1, \dots, n$ , then  $\ln L$  does not involve  $\alpha$  and is an increasing function of  $c$ . Hence,  $\hat{c} = +\infty$  and  $\hat{\alpha}$  is arbitrary. These estimates correspond to  $\hat{f}_0 = 1$  and  $\hat{f}_k = 0$ ,  $k > 0$ .

Case ii:  $c$  Known. For any fixed value of  $c$ , it can be shown that there exists at least one solution to  $\frac{\partial}{\partial \alpha} \ln L = 0$  for  $\alpha$  if at least one  $k_1 > 1$ . The solution may not be unique. If all  $k_1 = 0$ ,  $\ln L$  does not involve  $\alpha$ , so  $\hat{\alpha}$  is arbitrary. Finally, if all  $k_1$ 's = 1 or some  $k_1$ 's = 0 and the remaining  $k_1$ 's = 1, then there may be no finite value for  $\hat{\alpha}$ .

Case iii: Both  $\alpha$  and  $c$  Unknown. If at least one  $k_1 > 0$ , it is difficult, if not impossible, to show that unique MLEs always exist. However, it is possible to find the unique values of  $c$  which maximize (4.2.1) for fixed values of  $\alpha$  using the Newton-Raphson procedure for solving  $\frac{\partial}{\partial c} \ln L = 0$ . A line search on the variable  $\alpha$  can then be performed to obtain an approximate maximum for (4.2.1).

For the cases of type I and type II censoring, the MLEs for this failure model are obtained in the same manner as just described. For type I censoring after  $T$  cycles, the log-likelihood function is

$$\begin{aligned} \ln L = & \sum_{i=1}^r \ln[1 - \exp(-c)(k_{(i)} + 1)^\alpha] - c \sum_{i=1}^r \sum_{j=1}^{k_{(i)}} j^\alpha \\ & - c(n-r) \sum_{j=1}^T j^\alpha. \end{aligned} \quad (4.2.3)$$

For type II censoring with  $r+m$  failures at or before  $k_{(r)}$ ,  $m=0,1,\dots,n-r$ , the log-likelihood is

$$\begin{aligned} \ln L = & \sum_{i=1}^{r+m} \ln[1 - \exp(-c)(k_{(i)} + 1)^\alpha] - c \sum_{i=1}^{r+m} \sum_{j=1}^{k_{(i)}} j^\alpha \\ & - c(n-r-m) \sum_{j=1}^{k_{(r)}+1} j^\alpha. \end{aligned} \quad (4.2.4)$$

The derivatives of (4.2.3) and (4.2.4) have the same behavior as those of (4.2.1), so the MLEs may be obtained numerically as described for uncensored data.

The mean lifetime for this model is  $E(X) = \sum_{k=1}^{\infty} \exp(-c \sum_{j=1}^k j^\alpha)$ . By the results referenced in Section 2,  $E(X)$  exists for all  $c > 0$  and  $\alpha > -1$  and does not exist when  $\alpha < -1$  for any value of  $c > 0$ . For  $\alpha = -1$ ,  $E(X)$  exists when  $c > 1$  but does not exist for  $c \leq 1$ . When  $E(X)$  exists, it may be easily obtained to any desired degree of accuracy by summing an appropriate number of terms in the series. An estimate of  $E(X)$  can be found by using  $\hat{\alpha}$  and  $\hat{c}$ . For  $\alpha < 0$  (DFR), a lower bound for  $E(X)$  is  $\exp(-c)/[1 - \exp(-c)]$ , which can be estimated by replacing  $c$  with  $\hat{c}$ . For  $\alpha > 0$  (IFR), this quantity is an upper bound for  $E(X)$ .

## 5. EXAMPLES

Nelson [8] reported times to breakdown in minutes of an insulating fluid subjected to various test voltages. We took as an example the failure times at 36 kV and reduced each observation to the next lowest integer, which corresponds to observing the life test once per minute and recording the lifetime as "k" if the breakdown occurred in time interval  $[k, k+1)$ . Nelson [8] assumed

an underlying Weibull distribution for the continuous failure times. As examples, we will consider estimation from this data for the "discrete Weibull" model of Section 3, Models 2 and 3 of Section 4, and the nonparametric DFR model of Barlow, Bartholomew, Bremner, and Brunk [1].

The data that we use, as adapted from Nelson [8] for 36kV, are: 0,0,0,0,1,1,2,2,2,2,3,3,5,13,25 (minutes before failure). The computed MLEs of parameters, pmfs, Sfs, discrete hazards, and mean lifetimes for the various models are shown for comparison in Table 1. A hazard plot of these 15 observations indicated that the life distribution is DFR. This is supported by the MLE of  $\alpha$  in the "discrete Weibull hazard" model (Model 3) which was  $\hat{\alpha} \approx -.391$ .

It can be seen from Table 1 that the nonparametric DFR model yields the largest value for  $\ln L$ . This is to be expected since maximization of  $\ln L$  is done over all DFR distributions (including the other models in Table 1). The nonparametric model has two drawbacks, particularly when the number of failures is small. First, the estimate of the hazard function,  $\hat{h}_k$ , decreases at  $k$  only if there is an observed failure at  $k$ . Secondly, for  $k$  larger than the largest failure time,  $\hat{h}_k$  is not uniquely defined. For the purposes of Table 1  $\hat{h}_k \equiv \hat{h}_{25}$  for  $k > 25$ .

The three parametric models yield similar estimates of the failure distribution. We make the following distinctions among the models. In terms of  $\ln L$ , Model 2 performs the best for this data. Calculation of the estimators is more complicated for the discrete Weibull model than for Models 2 and 3. Model 3 and the discrete Weibull model are more flexible than Model 2 since they allow both IFR and DFR distributions. These proposed models provide a large class of distributions for fitting discrete failure data.

## ACKNOWLEDGMENTS

This research was supported by the United States Air Force Office of Scientific Research and Army Research Office under grant number AFOSR-84-0156.

## REFERENCES

- [1] R. E. Barlow, D. J. Bartholomew, J. M. Bremner, H. D. Brunk, Statistical Inference Under Order Restrictions, John Wiley & Sons, New York, 1972.
- [2] R. E. Barlow, F. Proschan, Statistical Theory of Reliability and Life Testing, Holt, Rinehart & Winston, New York, 1975.
- [3] Z. W. Birnbaum, S. C. Saunders, "Estimation for a family of life distributions with applications to fatigue," J. Applied Probability, vol 6; 1969 Jun, pp 328-347.
- [4] T. Fort, Infinite Series, Oxford University Press, London, 1930.
- [5] J. D. Kalbfleisch, R. L. Prentice, The Statistical Analysis of Failure Time Data, John Wiley & Sons, New York, 1980.
- [6] N. A. Langberg, R. V. León, J. Lynch, F. Proschan, "Extreme points of the class of discrete decreasing failure rate life distributions," Math. of Operations Research, vol 5, 1980 Feb, pp 35-42.
- [7] N. A. Langberg, R. V. León, J. Lynch, F. Proschan, "Extreme points of the class of discrete decreasing failure rate average life distributions," TIMS/Studies in the Management Sciences (S. H. Zanakis, J. S. Rustagi, eds), North-Holland Publishing Co., vol 19, 1982, pp 297-304.
- [8] W. Nelson, "Analysis of accelerated life test data - least squares methods for the inverse power law model," IEEE Transactions on Reliability, vol R-24, 1975, pp 103-107.
- [9] A. A. Salvia, R. C. Bollinger, "On discrete hazard functions," IEEE Transactions on Reliability, vol R-31, 1982 Dec, pp 458-459.

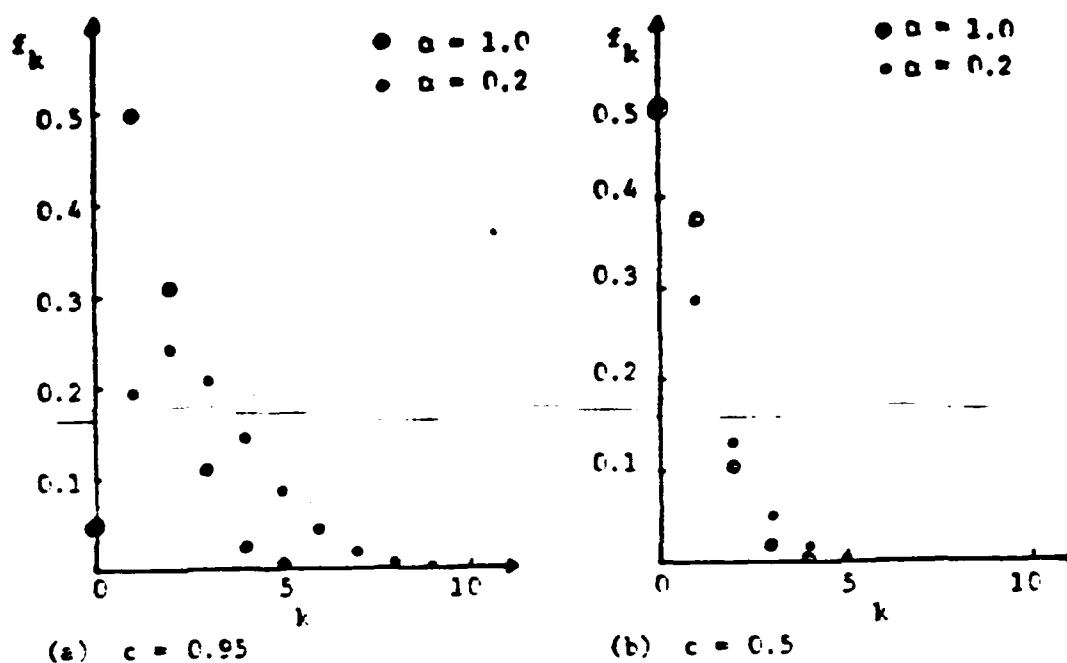


Figure 1: Model 1.

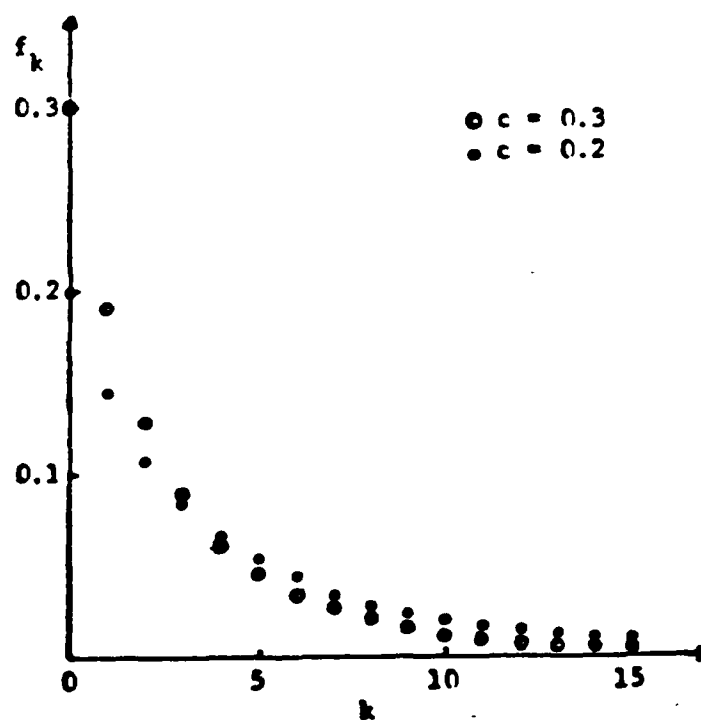
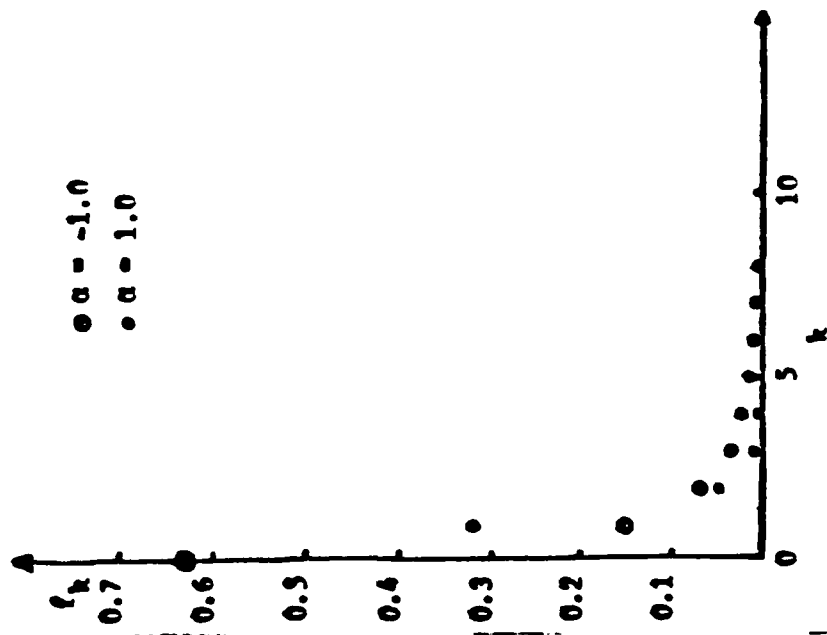
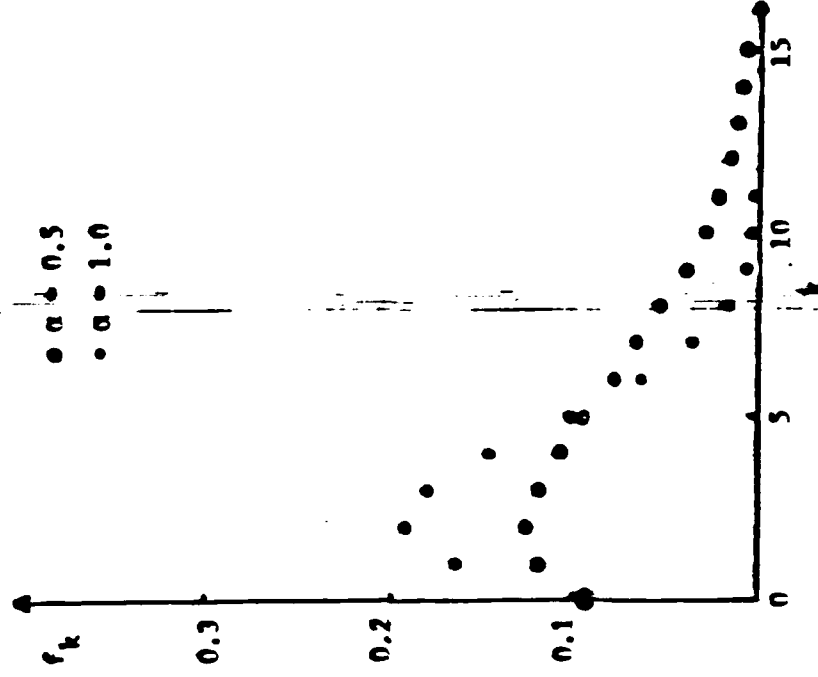


Figure 2: Model 2 with  $\alpha = 0.1$ .



(a)  $c = 1.0$



(b)  $c = 0.1$

Figure 3: Model 3 (Discrete Weibull Hazard Distribution).

### TABLE 1

## Estimated Failure Models

Empirical	Discrete Weibull				Model 2				Model 3				Nonparametric
	$\hat{\alpha}=.756$	$\hat{\beta}=.274$	$\hat{\alpha}=.127$	$\hat{c}=.298$	$\hat{\alpha}=-.391$	$\hat{c}=.375$							
	$-lnL = 36.324$				$-lnL = 35.922$				$-lnL = 36.095$				$-lnL = 34.343$
$\hat{E}(X)=3.933$	$\hat{E}(X)=3.865$				$\hat{E}(X)=4.092$				$\hat{E}(X)=3.894$				$\hat{E}(X)=4.307$
k	$\hat{f}_k$	$\hat{f}_k$	$\hat{h}_k$	$\hat{s}_k$	$\hat{f}_k$	$\hat{h}_k$	$\hat{s}_k$	$\hat{f}_k$	$\hat{h}_k$	$\hat{s}_k$	$\hat{f}_k$	$\hat{h}_k$	$\hat{s}_k$
0	.2667	.3133	.313	1.000	.2982	.298	1.000	.3130	.313	1.000	.3000	.300	1.000
1	.1333	.1566	.228	.687	.1858	.265	.702	.1710	.249	.687	.2100	.300	.700
2	.2667	.1079	.204	.530	.1228	.238	.516	.1118	.217	.516	.1470	.300	.490
3	.1333	.0798	.189	.422	.0850	.216	.393	.0793	.196	.404	.1029	.300	.343
4	.0000	.0612	.179	.342	.0610	.198	.308	.0589	.181	.325	.0400	.167	.240
5	.0667	.0481	.171	.281	.0451	.183	.247	.0452	.170	.266	.0333	.167	.200
6	.0000	.0384	.165	.233	.0342	.169	.202	.0355	.161	.221	.0119	.071	.167
7	.0000	.0311	.160	.195	.0265	.158	.168	.0284	.153	.185	.0111	.071	.155
8	.0000	.0254	.155	.164	.0209	.148	.141	.0231	.147	.157	.0103	.071	.144
9	.0000	.0209	.151	.138	.0168	.139	.120	.0189	.142	.134	.0095	.071	.134
11	.0000	.0145	.145	.100	.0112	.125	.090	.0131	.132	.099	.0082	.071	.115
13	.0667	.0103	.140	.073	.0078	.113	.069	.0939	.125	.075	.0071	.071	.099
15	.0000	.0074	.135	.054	.0057	.103	.055	.0069	.119	.058	.0061	.071	.086
20	.0000	.0034	.127	.027	.0028	.084	.033	.0034	.108	.031	.0042	.071	.059
25	.0667	.0017	.121	.014	.0016	.072	.022	.0018	.100	.018	.0029	.071	.041
35	.0000	.0004	.112	.004	.0006	.055	.012	.0006	.088	.007	.0014	.071	.019
50	.0000	.0001	.103	.001	.0002	.041	.006	.0001	.078	.002	.0005	.071	.006

**END**

**FILMED**

**10-84**

**DTIC**